Augmented Democracy as a Coherence-Constrained Control System

Governance Infrastructure Under Adversarial Conditions

Sylvain Cormier
Paraxiom Research
sylvain@paraxiom.org

December 2025

Abstract

We present a systems-theoretic framework for democratic governance that treats legitimacy as the preservation of coherence in decision-making processes rather than as a direct consequence of majority outcomes. The framework introduces *procedural infrastructure*—including token-curated test grids, dynamic credential NFTs, and bounded voting weights—that restrict participation to actors who can demonstrate verifiable engagement with relevant evidence, **without assigning semantic authority to the system itself**.

Building on prior work in Hamiltonian machine learning (ERLHS), geometric consensus (Karmonic Mesh), and quantum-safe validation (Proof of Coherence), we formalize augmented democracy as a control system in which proposals are state transitions, participants act as distributed sensors with bounded influence, and acceptance requires satisfaction of both a democratic majority condition and a coherence threshold measuring process consistency.

The central claim is that **democratic legitimacy is a measurable property of process quality—independent of specific outcomes**—and that governance systems designed around procedural admissibility and coherence constraints exhibit improved resistance to manipulation, Sybil attacks, plutocratic capture, and coordinated adversarial behavior.

Contents

1 Problem: Democracy Under Adversarial Load

Democratic systems face compounding failure modes in the information age. These failures are not political in nature—they are *systems failures* that require engineering solutions.

1.1 The Five Failure Modes

Unengaged Voting. Participants vote without verifiable engagement with proposal-relevant evidence. In traditional systems, nothing prevents a voter from casting a ballot on a complex policy question without having read a single document about it. The resulting signal is noise, not preference revelation.

Coordinated Manipulation. Influence campaigns exploit cognitive biases at scale. Social media amplification, bot networks, and targeted advertising can shift public opinion without any engagement with factual evidence. The democratic process becomes a contest of manipulation rather than deliberation.

Speed vs. Legitimacy Tradeoff. Digital communication enables rapid decision-making, but legitimate deliberation requires time. Systems optimized for speed sacrifice the careful consideration that gives decisions their legitimacy. The result is "fast wrong" instead of "slow right."

AI-Generated Influence. Large language models enable scalable persuasion. A single adversary can generate thousands of personalized, convincing arguments. The asymmetry between offense (generating influence) and defense (verifying authenticity) favors attackers.

Quantum-Era Threats. Future adversaries with quantum computers can retroactively compromise classical cryptographic commitments. Decisions made today with ECDSA signatures may be forgeable in a decade. Governance infrastructure must be quantum-resistant from the start.

1.2 Why Traditional Responses Fail

Traditional responses to these failures—education, regulation, fact-checking—address symptoms rather than architecture.

- **Education** assumes voters will become informed given sufficient resources. But rational ignorance [7] suggests voters have no individual incentive to invest in becoming informed.
- **Regulation** assumes platforms can be forced to reduce manipulation. But platform incentives align with engagement, not accuracy, and regulatory capture is endemic.
- **Fact-checking** assumes authoritative sources can determine truth. But "who fact-checks the fact-checkers?" leads to infinite regress, and institutional trust has collapsed across the political spectrum.

We require a different approach: treating democracy as *critical infrastructure* subject to engineering constraints, with formal guarantees that do not depend on participant virtue or institutional trustworthiness.

2 Philosophical Foundations: Artifacts, Not Truth

Before presenting the technical framework, we establish a critical philosophical distinction that underlies the entire system.

2.1 The Semantic Authority Problem

Any system that claims to determine "what is true" faces an infinite regress: who verifies the verifiers? Classical epistemic gatekeeping (expert panels, editorial boards, fact-checkers) ultimately rests on institutional authority that can be captured, corrupted, or contested.

Consider the problem concretely:

- If we say "experts determine facts," we must answer: which experts? Selected by whom? Accountable to whom?
- If we say "consensus determines facts," we must answer: consensus of whom? Over what time period? Weighted how?
- If we say "evidence determines facts," we must answer: evidence judged by whom? According to what standards?

Every answer leads to another question. There is no non-circular foundation for semantic authority.

2.2 The Solution: Procedural Authority

Our framework sidesteps this problem entirely. The system makes no semantic judgments. It enforces *procedural constraints* on what evidence may be referenced and measures *statistical properties* of the decision process.

This is not a weakness but a design requirement. A system that claimed semantic authority would be both philosophically indefensible and practically capturable.

2.3 Admissible Artifacts

Definition 1 (Admissible Fact Artifact)**.** *In this framework, a* fact *is not treated as a semantic claim or assertion, but as a verifiable artifact with cryptographic provenance. An artifact is admissible for use in governance processes if it satisfies:*

1. ***Immutable referenceability****: the artifact can be uniquely identified via a hash, DOI, or equivalent content-addressed reference;*
2. ***Authentic issuance****: the artifact is bound to a recognized issuer, authority, or origin via digital signatures or equivalent mechanisms;*
3. ***Contextual relevance****: the artifact belongs to an artifact class declared relevant for the proposal type under consideration.*

No semantic interpretation, truth evaluation, or correctness judgment is performed by the governance system itself.

This definition has important consequences:

- A peer-reviewed paper is admissible because it has a DOI, is signed by a journal, and belongs to the "scientific literature" artifact class—not because its conclusions are "true."
- A government report is admissible because it has a document ID, is signed by an agency, and belongs to the "official statistics" artifact class—not because its numbers are "correct."
- A dissenting scientific paper is equally admissible to a consensus paper, provided both have valid provenance. The system does not adjudicate which is "right."
- Participants demonstrate engagement with admissible artifacts, not agreement with their conclusions.

2.4 Test Grids as Admissibility Registries

Definition 2 (Token-Curated Test Grid). *A Token-Curated Test Grid (TCTG) is a governed registry that specifies which classes of admissible artifacts may be referenced for a given proposal domain. Curation decisions are made by participants who stake tokens and are subject to economic penalties if admitted artifacts are later shown to be malformed, inauthentic, or procedurally invalid.*

Test grids govern admissibility criteria *(format, provenance, relevance), not semantic conclusions or interpretations.*

Critical distinction: Test grids do not define truth; they define the set of artifacts that a decision process is permitted to reference.

Curators are slashed for admitting artifacts that:

- Have invalid provenance (forged signatures, broken hash references)
- Violate format requirements (wrong artifact class for proposal type)
- Were retracted or invalidated by their issuing authority

Curators are *not* slashed for admitting artifacts whose conclusions are later contested, revised, or overturned through normal scientific or institutional processes. The system does not adjudicate semantic disputes.

2.5 Coherence as Process Quality

Definition 3 (Process Coherence). *The coherence score γ does not measure correctness, factual truth, or proposal merit. It measures statistical consistency and entropy quality of the decision process given the declared admissible evidence and participation constraints.*

Low coherence indicates process instability, coordinated behavior, or insufficient entropy, independent of proposal content.

A proposal can have high coherence and still be "wrong" by external standards. A proposal can have low coherence despite being "correct." The coherence score measures whether the *process* that produced the decision exhibited properties consistent with legitimate deliberation.

2.6 Scope of System Authority

The governance system constrains *how* decisions are made, not *what* decisions must conclude.

System Does	**System Does Not**
Verify artifact provenance	Evaluate artifact truth
Enforce participation requirements	Determine correct outcomes
Measure process consistency	Judge proposal merit
Bound individual influence	Override collective judgment
Detect coordinated manipulation	Censor unpopular positions

This scope limitation is not a weakness but a design requirement. A system that claimed semantic authority would be both philosophically indefensible and practically capturable.

3 Augmented Democracy: Definition

3.1 What Augmented Democracy Is Not

Before defining what augmented democracy *is*, we clarify what it is *not*:

- **Not technocracy**: Experts do not override participants. The system ensures engagement with evidence, not deference to expert opinion.
- **Not algorithmic rule**: No AI makes final decisions. The algorithm enforces process constraints; humans make substantive choices.
- **Not censorship**: All proposals enter the system. The filtering occurs on *participation*, not *content*.
- **Not plutocracy**: Wealth does not determine influence. Quadratic costs and reputation caps bound the advantage of resources.
- **Not truth arbitration**: The system does not determine what is "true." It determines what is procedurally valid.

3.2 Formal Definition

Definition 4 (Augmented Democracy)**.** *A **constrained participation system** with admissibility gates, where:*

1. *Participation requires registration (identity binding via dynamic NFTs)*
2. *Voters must demonstrate engagement with proposal-relevant admissible artifacts*
3. *Proposals traverse a filtering pipeline (review period with gatherers/curators)*
4. *Votes are weighted by reputation and bounded by quadratic costs*
5. *Acceptance requires coherence threshold satisfaction (process quality)*
6. *All transitions produce cryptographic proofs (auditability)*
7. *Credentials decay without ongoing participation (life-sustaining NFTs)*

The key distinction from classical democracy: legitimacy derives from *process invariants*, not outcome ratification. The "augmented" refers to procedural augmentation—ensuring voters have verifiably engaged with relevant evidence before their input shapes collective decisions.

3.3 The Augmentation Hierarchy

Different engagement levels are supported:

Role	**Engagement**	**Responsibility**
Voter	Pass engagement verification	Reputation stake
Contributor	Add sources to proposals	Token compensation
Gatherer	Identify admissible artifacts	Quality incentives
Curator	Assemble test grids	Token stake (slashable)
Validator	Verify coherence	Block rewards

Users self-select into roles based on desired involvement level.

4 The Coherence Pipeline

The augmented democracy system operates as a six-stage pipeline where each stage gates the next. This is not a collection of mechanisms but a *processing pipeline* with formal invariants at each transition.

4.1 Stage 1: Proposal Creation

A registered `Submitter` creates a proposal. The submitter must hold a valid credential NFT with the `Submitter` role. Proposals enter the system in `Submitted` status.

Gate: Valid submitter credential

On failure: Proposal rejected immediately

4.2 Stage 2: Review Period

Proposals enter a mandatory review period. During review:

- Reviewers assess proposal clarity and scope
- Duplicate or malformed proposals are flagged
- Community can signal concerns

Proposals that survive the review period advance to `ReadyForVoting`.

Gate: Review period completes without rejection

On failure: Proposal held for revision

Coherence dimension: Temporal coherence—deliberation has minimum dwell time.

4.3 Stage 3: Grid Curation

Parallel to review, credentialed **Gatherers** and **Curators** build the test grid:

1. Gatherers identify admissible artifacts: documents with valid provenance from recognized issuers (journals, agencies, standards bodies)
2. Curators assemble artifacts into engagement verification grids
3. Curators stake tokens on grid quality
4. Grid is published and linked to the proposal

Gate: Curator stake sufficient, grid properly formed

On failure: Grid flagged, new curators assigned

Coherence dimension: Procedural coherence—the artifact registry is curated by accountable participants with skin in the game.

4.4 Stage 4: Engagement Verification

Before voting, each participant must demonstrate engagement with proposal-relevant artifacts.

The engagement verification:

- Does not measure intelligence or political alignment
- Verifies that the voter has *encountered* the admissible artifacts
- Allows disagreement with artifact conclusions

Critical clarification: A voter who has engaged with the evidence and *disagrees* with its conclusions still passes the gate. The system verifies engagement, not agreement.

Gate: Engagement threshold met

On failure: Voter ineligible for this proposal; may retry after review

4.5 Stage 5: Weighted Voting

Eligible voters cast weighted votes. Vote weight is a function of:

1. **Reputation**: Accumulated from contributions, engagement verifications, prior participation
2. **Quadratic cost**: Multiple votes on same proposal cost n^2
3. **Quantum entropy**: $\pm 10\%$ randomization prevents prediction

Gate: Signature valid, not a duplicate

On failure: Vote rejected

4.6 Stage 6: Coherence Check and Result

After the voting period, finalization computes the result. Two conditions must **both** be satisfied:

1. **Democratic condition**: $w_{\text{approve}} > w_{\text{reject}}$
2. **Coherence condition**: $\gamma > 50$

Gate: Both conditions satisfied

On failure: No state change; may investigate and revote

4.7 Why Both Gates Are Necessary

Attack	Blocked by Engagement	Blocked by γ
Unengaged voting	✓	—
Sybil (fake identities)	Partial	✓
Coordinated manipulation	—	✓
Bribery	—	✓
Credential theft	—	✓
Bot voting	✓	✓

- **Admissibility gate**: Ensures voters have *encountered* relevant artifacts
- **Coherence gate** (γ): Ensures votes are *statistically independent*

A voter who demonstrates engagement but coordinates with others will trigger low γ. A voter who votes independently but hasn't engaged fails the admissibility gate. Both gates must pass for legitimate outcomes.

5 Procedural Infrastructure

This section describes the infrastructure components that implement the coherence pipeline.

5.1 Token-Curated Test Grids

Following the KILT Protocol model [5], test grids employ three distinct roles:

1. **Claimer**: A participant who asserts readiness to vote on a proposal
2. **Attester**: A curator who has assembled the test grid for that proposal
3. **Verifier**: The system that checks engagement verification completion

This separation prevents conflicts of interest: those who create tests do not administer them, and those who verify do not profit from outcomes.

5.1.1 Gatherers and Curators

Within the Attester role, two sub-functions operate:

- **Gatherers**: Identify admissible artifacts for a proposal—documents with valid provenance from recognized issuers. Gatherers are compensated per artifact that meets admissibility criteria.
- **Curators**: Assemble gathered artifacts into engagement verification grids. The goal is to confirm that voters have *encountered* the relevant evidence, not that they agree with it. Curators stake tokens on grid quality.

5.1.2 Economic Incentives and Slashing

Slashing conditions (curators lose staked tokens):

- Admitting artifacts with invalid provenance (forged signatures, broken hashes)
- Admitting artifacts from non-recognized issuers
- Admitting artifacts outside the declared artifact class
- Admitting artifacts that were retracted by their issuing authority

Not slashable (curators are protected):

- Admitting artifacts whose conclusions are later contested or revised
- Admitting artifacts that some participants disagree with
- Admitting artifacts from one scientific position when others exist

The system does not adjudicate semantic disputes. Curators are accountable for *procedural validity*, not *correctness*.

5.2 Dynamic Credential NFTs

Participant credentials are not static. A voter's eligibility, weight, and privileges evolve based on contribution history through **Dynamic NFTs**—non-fungible tokens whose metadata updates based on on-chain activity.

The credential NFT updates automatically when:

- An engagement verification is passed or failed
- A vote is cast
- A proposal is submitted
- Reputation is adjusted by peer review
- Domain certification is earned or revoked

5.2.1 Life-Sustaining NFTs

A critical innovation is the **Life-Sustaining NFT**—a credential that requires ongoing activity to remain valid. Unlike static credentials that persist indefinitely, life-sustaining NFTs decay without continuous participation.

This addresses the "ghost voter" problem: credentials issued to participants who subsequently disengage. By requiring periodic activity (voting, contributing, engagement verification), the system ensures that voting weight reflects *active* participation, not historical registration.

5.2.2 Domain-Specific Credentials

Voters may hold credentials in specific domains:

Domain	Unlocks Voting On
Environmental	Climate, conservation, pollution proposals
Technical	Infrastructure, protocol upgrades
Economic	Treasury, tokenomics, funding proposals
Social	Community guidelines, dispute resolution
Emergency	Crisis response, security incidents

Domain credentials are earned by passing domain-specific engagement verifications and maintained through ongoing participation.

5.3 Quadratic Voting for Bounded Influence

The framework incorporates **quadratic voting** to prevent plutocratic capture while preserving signal strength for high-conviction preferences.

5.3.1 The Quadratic Cost Function

The cost to cast n votes on a single proposal from a single participant:

$$\text{cost}(n) = n^2 \tag{1}$$

Votes Cast	Cost	Marginal Cost
1	1	1
2	4	3
3	9	5
4	16	7
5	25	9

The increasing marginal cost discourages concentration of voting power on single proposals, encouraging participants to distribute influence across multiple issues.

5.3.2 Whale Resistance

Consider an adversary with 100$\times$ the resources of an average participant:

System	Adversary Influence	Ratio
1-person-1-vote	1 vote	1:1
Plutocratic (1:1 stake)	100 votes	100:1
Quadratic	10 votes	10:1
Quadratic + Reputation Cap	$\leq$ 10 votes	$\leq$ 10:1

Quadratic voting reduces the 100:1 wealth advantage to a 10:1 voting advantage. Combined with reputation caps and coherence thresholds, adversarial influence is further bounded.

6 Governance as a Control System

Democratic governance can be formally modeled as a control system with specific invariants that must be preserved across state transitions.

6.1 The Control-Theoretic View

Definition 5 (Governance Control System)**.** *A tuple* (S, A, T, I, R) *where:*

- S*: State space (current policy/system configuration)*
- A*: Action space (proposals as potential state transitions)*
- $T : S \times A \to S$*: Transition function (proposal execution)*
- I*: Coherence invariant (must hold for valid transitions)*
- R*: Rejection mechanism (blocks invalid transitions)*

6.2 Participants as Sensors

In control theory, sensors measure system state. In governance, participants provide distributed measurement of collective preference. Unlike physical sensors, governance sensors are:

- **Heterogeneous**: Each participant has different information access
- **Strategic**: Participants may misreport to influence outcomes
- **Bounded**: Each participant sees only part of the system

The framework addresses these properties through credential-weighted aggregation and coherence-based filtering.

6.3 The Dual Condition

The central contribution is the dual-condition consensus requirement. Approval requires **both**:

1. $w_{\text{approve}} > w_{\text{reject}}$ (democratic condition)
2. $\gamma > 50$ (coherence condition)

where γ is the *quantum confidence score*, computed as:

$$\gamma = \min\left(\frac{\sigma_\eta^2}{255}, 1\right) \times 100 \tag{2}$$

and σ_η^2 is the variance of entropy values across votes.

Key insight: Low variance indicates correlated votes (Sybil attack, coordination, bribery). High variance indicates independent voting (legitimate process).

6.4 Coherence as Lyapunov Function

The coherence threshold acts as a Lyapunov function. States with $\gamma \leq 50$ are *outside the coherence manifold*—they may satisfy the democratic condition but fail the process quality invariant.

The system refuses to transition to such states, even under majority pressure. This inverts the traditional democratic assumption: legitimacy is not derived from majority agreement alone, but from the quality of the process that produced the agreement.

6.5 Bounded Influence

No single participant can dominate outcomes. The weight function satisfies:

$$w_i \in [w_{\min}, w_{\max}] \quad \forall i \tag{3}$$

with $w_{\min} = 1$ (floor) and $w_{\max} = r_i \cdot 1.1$ (reputation cap with 10% entropy bonus maximum).

6.6 Rollback and Retry Semantics

Failed coherence checks do not corrupt state. The system remains at s_t when:

- $w_{\text{approve}} \leq w_{\text{reject}}$ (democratic failure)
- $\gamma \leq 50$ (coherence failure)
- Entropy pool exhausted (resource failure)

Submitters may revise and resubmit. The state machine supports retry without penalty.

7 Coherence as Layered Invariant

The term *coherence* appears across multiple domains in distributed systems, from quantum physics to machine learning to governance. This section clarifies how procedural coherence (γ) relates to these other usages and argues that coherence preservation is a unifying principle across the full system stack.

7.1 Three Layers of Coherence

We identify three distinct but related coherence concepts:

Definition 6 (Quantum Coherence)**.** *In quantum information systems, coherence refers to the preservation of phase relationships between quantum states. A quantum channel maintains coherence when superposition states survive transmission without collapsing into classical mixtures. Formally, for density matrix ρ:*

$$\mathcal{C}_q(\rho) = \sum_{i \neq j} |\rho_{ij}|$$

measures the off-diagonal elements that encode quantum correlations. Decoherence—the loss of these correlations through environmental interaction—degrades channel capacity and security guarantees in protocols like QKD.

Definition 7 (Geometric Coherence)**.** *In representation learning and neural architectures, coherence refers to the preservation of semantic relationships in embedding spaces. A model maintains geometric coherence when:*

1. *Similar inputs map to proximate regions of the embedding manifold*

2. *Semantic operations (analogy, composition) correspond to geometric operations*

3. *The manifold structure remains stable under perturbation*

Geometric incoherence manifests as embedding drift, manifold collapse, or semantic inconsistency—conditions where the model's internal representations no longer faithfully encode the relationships they purport to represent.

Definition 8 (Procedural Coherence)**.** *In governance systems, coherence refers to the statistical properties of decision processes that indicate authentic, independent participation. As defined in Section 4, the procedural coherence score:*

$$\gamma = \frac{\sigma_\eta^2}{255} \times 100$$

measures entropy-weighted variance in voting patterns. Low γ indicates coordinated behavior inconsistent with independent decision-making.

7.2 The Coherence Stack

These three coherence types form a dependency hierarchy in systems that span physical infrastructure through algorithmic mediation to collective decision-making:

Physical → Model: The entropy injection mechanism ($\epsilon \sim \pm 10\%$) that enables procedural coherence detection relies on high-quality randomness. If quantum entropy sources suffer decoherence, the resulting pseudo-randomness may exhibit patterns that either (a) fail to distinguish authentic from coordinated voting, or (b) introduce false positives by injecting correlated noise. Quantum coherence in the entropy generation layer is thus a precondition for meaningful γ measurement.

Model → Governance: If test grid verification or artifact validation employs machine learning models (e.g., for semantic relevance checking, engagement verification, or anomaly detection), geometric coherence in those models affects governance outcomes. A model with collapsed embeddings may fail to distinguish substantively different proposals; a model with drifted representations may inconsistently apply admissibility criteria. Geometric coherence in the model layer is thus a precondition for consistent procedural application.

7.3 Coherence Degradation Propagates Upward

A critical property of the coherence stack is that degradation propagates upward but not downward:

Proposition 1 (Upward Propagation)**.** *Let $\mathcal{C}_q$, $\mathcal{C}_g$, and γ denote quantum, geometric, and procedural coherence respectively. Then:*

$$\mathcal{C}_q < \mathcal{C}_q^{\min} \implies \mathbb{E}[\gamma] < \gamma_{authentic}$$

and

$$\mathcal{C}_g < \mathcal{C}_g^{\min} \implies Var[\gamma] > Var_{expected}$$

That is, quantum decoherence biases procedural coherence measurements, and geometric incoherence increases their variance.

This has practical implications: a governance system cannot achieve reliable procedural coherence guarantees without monitoring coherence at lower layers. The $\gamma > 50$ threshold assumes baseline coherence in entropy sources and verification models.

7.4 Unified Interpretation

Despite domain-specific formalizations, all three coherence types share a common interpretation:

> *Coherence measures the degree to which a system preserves the structure it claims to preserve.*

- **Quantum**: Preserves superposition structure through transmission
- **Geometric**: Preserves semantic structure through embedding
- **Procedural**: Preserves independence structure through aggregation

In each case, coherence loss indicates that the system's outputs no longer faithfully represent its inputs according to the intended transformation. A decoherent quantum channel does not faithfully transmit quantum states; an incoherent embedding does not faithfully represent semantic relationships; an incoherent vote does not faithfully aggregate independent preferences.

7.5 Implications for System Design

The layered coherence model suggests several design principles:

1. **Monitor all layers**: Governance coherence (γ) should be accompanied by monitoring of entropy source quality and model stability. Anomalies at lower layers may explain or predict governance-layer anomalies.
2. **Establish layer-specific thresholds**: Just as $\gamma_{\min} = 50$ gates governance transitions, analogous thresholds should gate reliance on entropy sources ($\mathcal{C}_q^{\min}$) and verification models ($\mathcal{C}_g^{\min}$).
3. **Fail safely on coherence loss**: When lower-layer coherence degrades, the system should either (a) halt governance operations, (b) fall back to higher-threshold requirements, or (c) switch to backup sources with intact coherence.
4. **Audit coherence dependencies**: System audits should explicitly trace which coherence guarantees depend on which lower-layer assumptions, enabling targeted hardening.

7.6 Scope Clarification

This paper focuses on procedural coherence (γ) and its role in governance integrity. The connections to quantum and geometric coherence are noted to:

1. Clarify terminology for readers familiar with those domains
2. Identify dependencies that affect γ reliability
3. Suggest a broader research program on coherence-preserving systems

Full treatment of quantum coherence in entropy generation and geometric coherence in verification models is deferred to companion work. The procedural coherence results in this paper hold under the assumption that lower-layer coherence is maintained above implementation-specific thresholds.

8 Threat Model and Defenses

8.1 What γ Detects

The coherence score γ detects attacks that produce correlated vote patterns:

- **Sybil attacks**: Fake identities controlled by a single adversary produce correlated entropy signatures, resulting in low variance.
- **Bribery**: Vote buying produces coordinated voting patterns, detectable as low entropy variance.
- **Coordination attacks**: Any form of lockstep voting—whether from a political party, interest group, or bot network—produces statistical anomalies that trigger low γ.
- **Replay attacks**: Reused entropy values produce zero variance, immediately failing the coherence check.

8.2 What γ Does Not Detect

The coherence score is explicitly *not* a measure of:

- Whether the proposal is "good" or "bad"
- Whether voters are "correct" in their positions
- Whether the outcome is desirable by any external standard
- Whether the proposal will have positive consequences

γ measures process quality, not outcome quality. A high-coherence vote can produce a "bad" outcome; a low-coherence vote might have produced a "good" outcome. The system makes no judgment.

8.3 Defense Summary

Attack	**Primary Defense**	**Detection Method**
Sybil	γ	Correlated entropy
Bribery	γ	Correlated voting patterns
Coordination	γ	Statistical anomaly
Unengaged voting	Engagement gate	Verification failure
Bot voting	Both	Pattern + engagement
Whale domination	Quadratic costs	n^2 scaling
Replay	Signature gate	Nonce validation

9 Historical Evolution

The augmented democracy framework has evolved through eight years of development, from theoretical conception to production deployment.

9.1 Conceptual Foundation (2017)

The original framework was published in August 2017 as "A Machine-Based Societal Model for Curbing Citizen Cynicism." Key concepts established:

- **Test grids**: Voters must demonstrate engagement with relevant evidence
- **Mechanical humans**: Low-barrier contribution tasks for universal participation
- **Operating system metaphor**: Governance as a self-regulating system
- **Blockchain validation**: Immutable record of all decisions
- **Official as safeguard**: Human oversight of machine recommendations

9.2 Implementation Timeline

Principle	2017	EOS	Substrate	QH
Artifact-gated voting	Concept	Implemented	Enhanced	Production
Dynamic credentials	Concept	Prototype	NFT-based	Life-sustaining
Quadratic bounds	Concept	Implemented	Integrated	+ Entropy
Coherence threshold	Implicit	Partial	Formal	Quantum
Official safeguard	Central	Preserved	Root-only	Emergency

2020–21: EOS smart contract prototype validated the concept
2022–23: Substrate pallet migration enabled formal verification
2024–25: Quantum Harmony deployment with post-quantum signatures

10 Failure Modes and Recovery

10.1 When Coherence Fails

Coherence failure ($\gamma \leq 50$) indicates:

1. **Coordinated attack**: Sybil voters with correlated entropy
2. **Entropy exhaustion**: Insufficient quantum randomness available
3. **System compromise**: Entropy source manipulation

Critically, coherence failure does *not* indicate that the proposal is wrong, harmful, or should be rejected on substantive grounds. It indicates only that the *process* exhibited statistical anomalies.

10.2 Recovery Mechanisms

Event	Detection	Response
Low entropy	Immediate	Auto-pause voting
Coherence failure	At finalization	Proposal held
Repeated failures	3 consecutive	Escalate to committee
System compromise	Manual detection	Emergency halt

The system includes a "break glass" mechanism for emergency governance, but root authority *cannot* directly approve or reject proposals—only pause the system and trigger investigation.

11 Conclusion: Democracy as Infrastructure

Democracy is no longer a philosophical abstraction. It is **critical infrastructure**.

The systems we use to make collective decisions are under adversarial load from state actors, AI-generated influence, and the fundamental speed/quality tradeoff of digital communication. Traditional democratic theory offers no defense against these threats because it treats legitimacy as a property of outcomes rather than processes.

This paper has presented an alternative: *coherence-constrained democratic systems* with *procedural infrastructure*, where:

1. Proposals are state transitions in a formal control system
2. Participants hold dynamic credentials that evolve with contribution
3. Voters must demonstrate engagement with admissible artifacts
4. Quadratic voting costs bound plutocratic influence
5. Deliberation is a filtering operation with minimum dwell time
6. Approval requires both majority support *and* coherence threshold
7. All transitions produce cryptographic proofs
8. Credentials decay without ongoing participation

The central thesis:

> **Democratic legitimacy is a measurable property of process quality, independent of specific outcomes.**

This is not ideology. It is engineering.

The governance system constrains how decisions are made, not what decisions must conclude.

Combined with prior work on ERLHS (coherence in AI) [2], Karmonic Mesh (geometric substrate) [3], and Proof of Coherence (distributed validation) [4], this framework provides a complete architecture for augmented democracy as 21st-century infrastructure.

References

[1] Cormier, S. (2017). *A Machine-Based Societal Model for Curbing Citizen Cynicism.* Unpublished manuscript.

[2] Cormier, S. (2025). *ERLHS: A Hamiltonian Framework for Coherence-Preserving Machine Intelligence.* Zenodo. DOI: 10.5281/zenodo.17928909

[3] Cormier, S. (2025). *Karmonic Mesh: Spectral Consensus on Toroidal Manifolds.* Zenodo. DOI: 10.5281/zenodo.17928991

[4] Cormier, S. (2025). *Proof of Coherence: QKD-Based Distributed Consensus.* Zenodo. DOI: 10.5281/zenodo.17929054

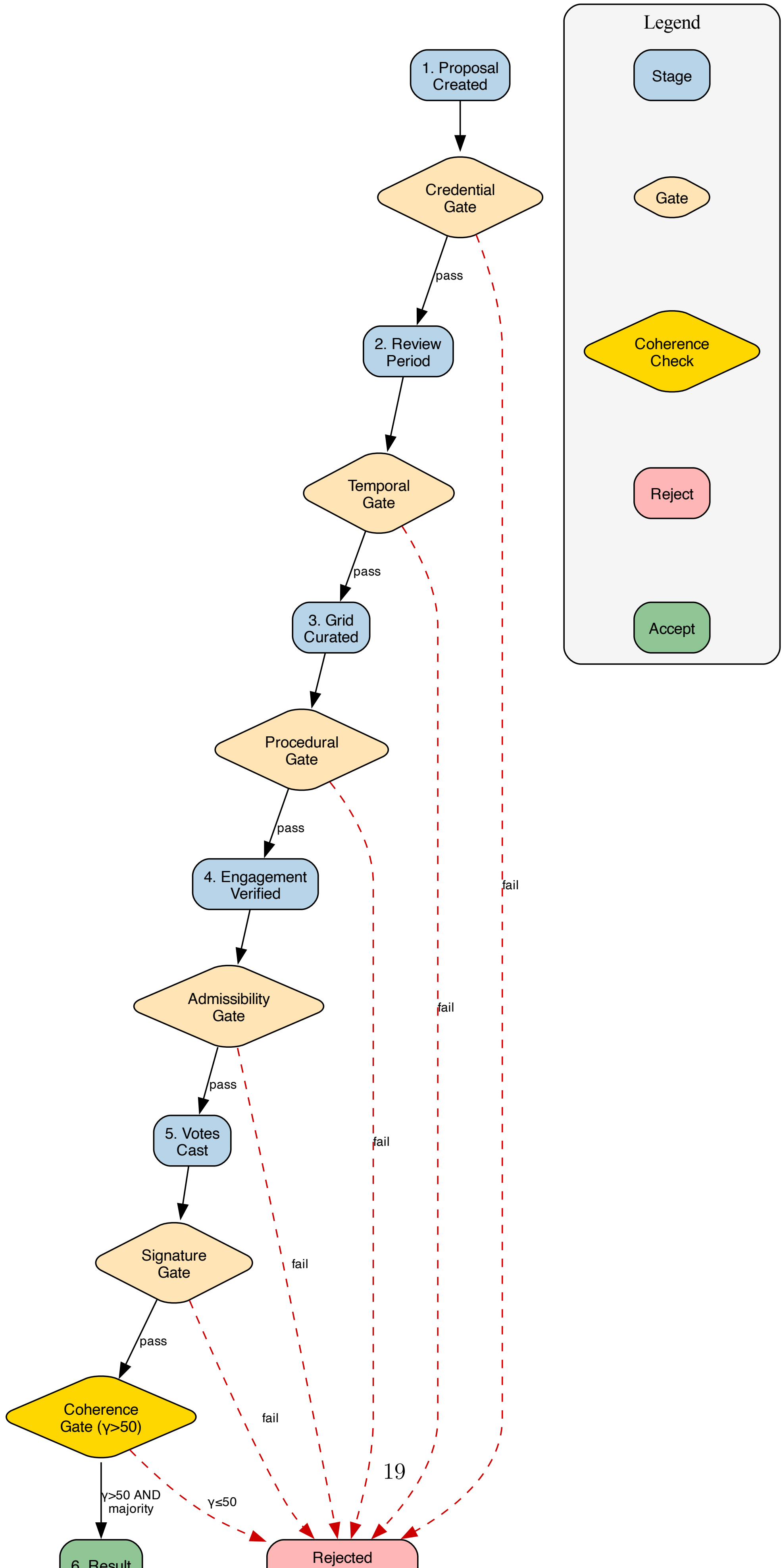
[5] BOTLabs GmbH. (2020). *KILT Protocol White Paper: Credentials for Web 3.0.* Berlin, Germany.

[6] Buterin, V., Hitzig, Z., & Weyl, E. G. (2019). *A Flexible Design for Funding Public Goods.* Management Science, 65(11), 5171–5187.

[7] Downs, A. (1957). *An Economic Theory of Democracy.* Harper & Row.

Technical Appendices

Formal definitions, protocol mechanics, security analysis, and implementation evidence are available in the companion document: *Augmented Democracy: Technical Appendices* (paper-appendix.pdf).



Legend
Stage
Gate
Coherence Check
Reject
Accept
1. Proposal Created
Credential Gate
pass
2. Review Period
Temporal Gate
pass
3. Grid Curated
Procedural Gate
pass
4. Engagement Verified
Admissibility Gate
pass
5. Votes Cast
Signature Gate
pass
Coherence Gate (γ>50)
γ>50 AND majority
γ≤50
fail
fail
fail
fail
fail
6. Result
Rejected

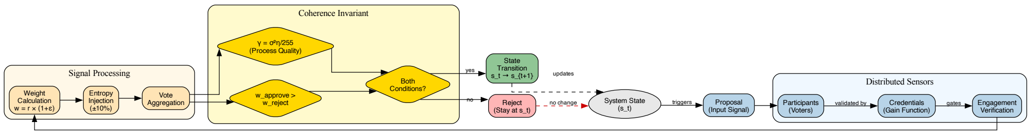


Figure 2: Governance as a Control Loop. Participants act as distributed sensors; credentials provide gain functions; the coherence check ($\gamma > 50$) gates state transitions.

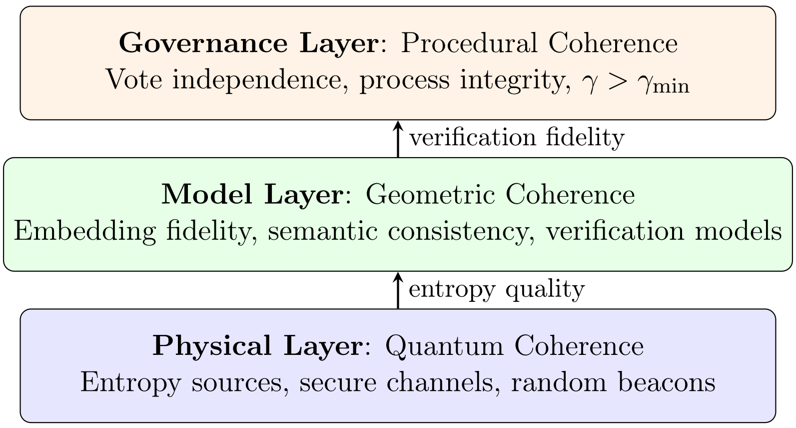


Figure 3: The coherence stack: each layer depends on coherence preservation in layers below.

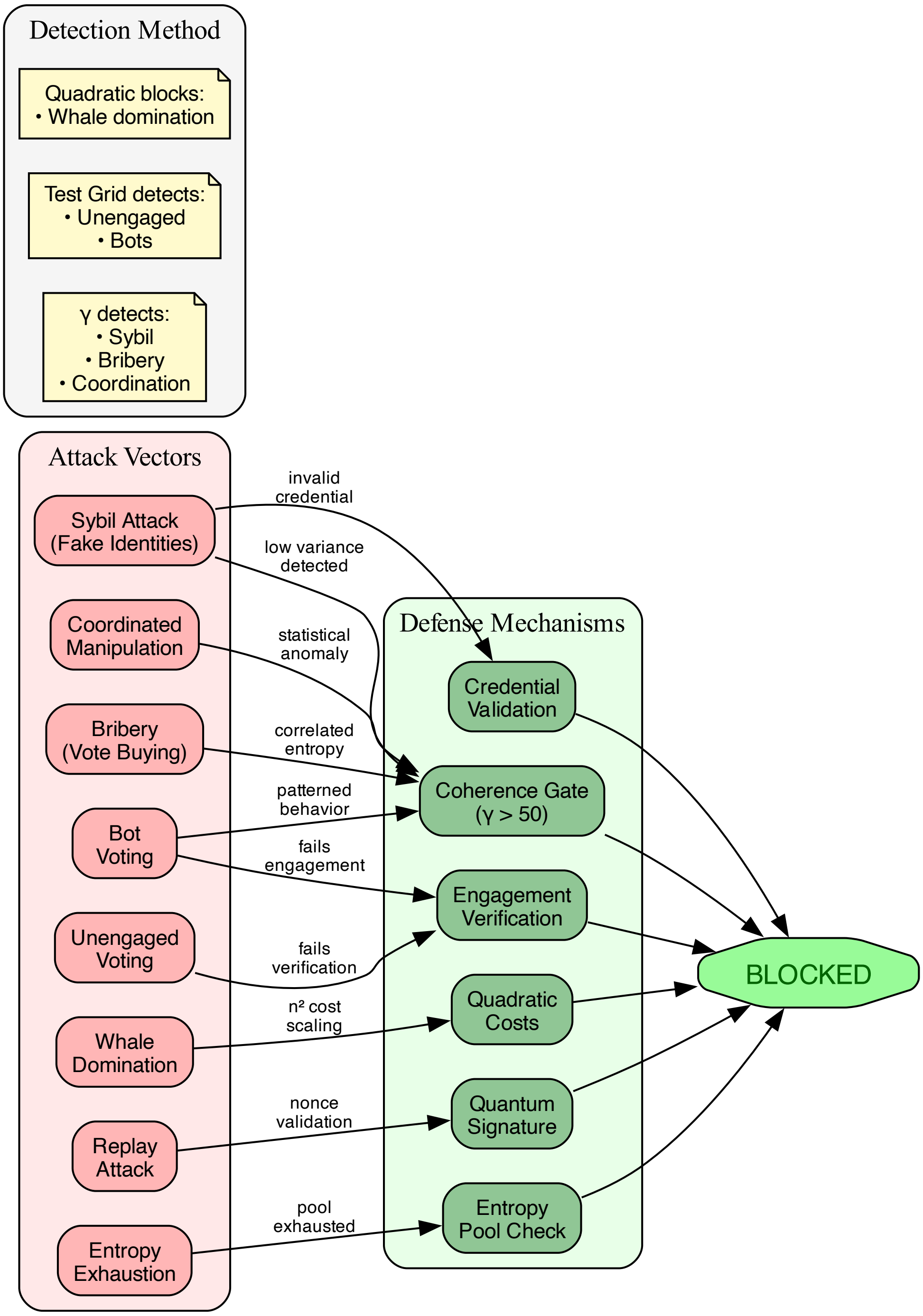


Figure 4: Attack vectors and the defense mechanisms that block them. The coherence gate (γ) detects coordination-based attacks; the engagement gate detects unengaged voting; quadratic costs bound whale domination.